

Correlation Properties of Quantum Measurements

PAUL BUSCH¹ AND PEKKA J. LAHTI²

ABSTRACT

The kind of information provided by a measurement is determined in terms of the correlation established between observables of the apparatus and the measured system. Using the framework of quantum measurement theory, necessary and sufficient conditions for a measurement interaction to produce strong correlations are given and are found to be related to properties of the final object and apparatus states. These general results are illustrated with reference to the *standard model* of the quantum theory of measurement.

PACS number: 03.65.Bz.

¹Department of Applied Mathematics, The University of Hull, Hull HU6 7RX, UK.

E-mail: p.busch@maths.hull.ac.uk

²Department of Physics, University of Turku, 20014 Turku, Finland.

E-mail: pekka.lahti@utu.fi

I. Introduction.

Any physical measurement is carried out in order to provide information about a specified system, its state prior to or after the measurement. The procedure generally is to ascertain the values of a pointer observable, which have become correlated with some observable of the measured system. Thus the kind of information available by a given measurement depends on the statistical dependencies established by the interaction between the apparatus, or some probe system, and the object.

The minimal content of the notion of measurement in quantum mechanics is given by the *probability reproducibility* requirement; according to this condition, a particular measurement scheme qualifies as a measurement of a given observable E if for all initial states of the object system the associated probability distributions of E are reproduced in the resulting statistics of pointer readings.¹ Regarding a large ensemble of object plus apparatus systems as one individual system, this situation can be described in terms of strong correlations between the values of the frequency operators [see, e.g., ref. 1] associated with the observable E and the pointer observable, respectively. In the present context we shall not be concerned with the ensembles regarded as individual objects but rather we shall analyze statistical dependencies between individual members of the ensembles as they show up in certain correlation quantities. The following three kinds of correlations appear naturally in the measurement context: correlations between an object observable E and the pointer observable; correlations between the values of these observables; and correlations between the conditional final states of the object system and apparatus, respectively. Our goal is to give exhaustive characterizations of the conditions under which these correlations are established. It will be found that the final component states of the object and apparatus must possess certain properties in order that such correlations may be strong.

Our investigation builds on previous work published in ref. 2. Correcting an erroneous characterization of strong correlations used in that paper, we give here a complete account of necessary and sufficient conditions for the occurrence of strong correlations. In addition, the scope of the results is extended beyond unitary measurements and sharp observables to cover arbitrary measurements and pointer observables and general object observables. Finally, possible fields of applications are indicated on the basis of the standard model of measurement theory, which was recently used in various proposals for quantum and atomic optics QND measurements.³

II. Framework.

2.1. We follow here the Hilbert space formulation of quantum mechanics in which the description of a physical system \mathcal{S} is based on a complex separable Hilbert space \mathcal{H} , with the inner product $\langle \cdot | \cdot \rangle$, and which builds on the dual concepts of states and observables reflecting the general structure of an experiment: preparation of the system followed by a measurement.

2.2. Let $\mathcal{L}(\mathcal{H})$ denote the set of bounded linear operators on \mathcal{H} and $\mathcal{T}(\mathcal{H})$ its subset of trace class operators. A *state* of \mathcal{S} is given as a positive linear operator

$T : \mathcal{H} \rightarrow \mathcal{H}$ of trace one. We let $\mathcal{S}(\mathcal{H}) := \{T \in \mathcal{T}(\mathcal{H}) \mid T \geq O, \operatorname{tr}[T] = 1\}$ denote the set of states of \mathcal{S} , and we recall that $\mathcal{S}(\mathcal{H})$ is a (σ) -convex subset of $\mathcal{T}(\mathcal{H})$, the one-dimensional projection operators $P[\varphi]$ (generated by the unit vectors $\varphi \in \mathcal{H}$) being its extremal elements. The $P[\varphi]$ shall be called *vector states*.

2.3. Let Ω be a set and \mathcal{F} a σ -algebra of subsets of Ω . An *observable* of \mathcal{S} is represented as (and identified with) a normalised positive operator valued measure, a POV measure, $E : \mathcal{F} \rightarrow \mathcal{L}(\mathcal{H})$, that is, an operator valued mapping $X \mapsto E(X)$ on \mathcal{F} with the properties: *i*) $E(\Omega) = I$, *ii*) $E(X) \geq O$, and $E(\cup X_i) = \sum E(X_i)$ for any disjoint sequence $(X_i) \subset \mathcal{F}$, with the series $\sum E(X_i)$ converging in the weak operator topology of $\mathcal{L}(\mathcal{H})$. We recall that the projection valued measures, the PV measures, are a particular case of the POV measures; furthermore a POV measure E is a PV measure, that is, $E(X)^2 = E(X)$ for all $X \in \mathcal{F}$, if and only if E is multiplicative, that is, $E(X \cap Y) = E(X)E(Y)$ for all $X, Y \in \mathcal{F}$. Observables which are represented by PV measure are called *sharp* observables. It is by now well-established that the extension of the notion of observables from PV measures to POV measures is a necessity in quantum mechanics.

2.4. The probability measure

$$(1) \quad p_T^E : \mathcal{F} \rightarrow [0, 1], \quad X \mapsto p_T^E(X) := \operatorname{tr}[TE(X)]$$

defined by an observable E and a state T is related to a measurement by virtue of the minimal interpretation of quantum mechanics: *the number $p_T^E(X)$ is the probability that a measurement of the observable E performed on the system \mathcal{S} in the state T leads to a result in the set X .* The intended empirical content of this statement is the following: if the same E -measurement were repeated sufficiently many times under the same conditions (characterised by T), then in the long run the relative frequency of the occurrence of the measurement results in X would approach the number $p_T^E(X)$.

III. Measurement.

A. General.

3.1. A *measurement scheme* for the (object) system \mathcal{S} consists of a *measuring apparatus* \mathcal{A} [with its Hilbert space $\mathcal{H}_{\mathcal{A}}$], a *pointer observable* $Z : \mathcal{F}_{\mathcal{A}} \rightarrow \mathcal{L}(\mathcal{H}_{\mathcal{A}})$ [with its ‘space of values’ $(\Omega_{\mathcal{A}}, \mathcal{F}_{\mathcal{A}})$], an *initial state* $T_{\mathcal{A}}$ of the apparatus, a *measurement coupling* V [a linear state transformation on $\mathcal{T}(\mathcal{H} \otimes \mathcal{H}_{\mathcal{A}})$], and a [measurable] *pointer function* $f : \Omega_{\mathcal{A}} \rightarrow \Omega$, with the assumption that if $T \in \mathcal{S}(\mathcal{H})$ is an initial state of \mathcal{S} , then $V(T \otimes T_{\mathcal{A}})$ is the final state of the compound object-apparatus system $\mathcal{S} + \mathcal{A}$. Taking the partial traces of $V(T \otimes T_{\mathcal{A}})$ over $\mathcal{H}_{\mathcal{A}}$ and \mathcal{H} , respectively, one gets the corresponding reduced states $\mathcal{R}_{\mathcal{S}}(V(T \otimes T_{\mathcal{A}}))$ and $\mathcal{R}_{\mathcal{A}}(V(T \otimes T_{\mathcal{A}}))$ of \mathcal{S} and \mathcal{A} , respectively; then the probability measure of the pointer observable Z in the final apparatus state is completely determined as $Y \mapsto p_{\mathcal{R}_{\mathcal{A}}(V(T \otimes T_{\mathcal{A}}))}^Z(Y) = \operatorname{tr}[\mathcal{R}_{\mathcal{A}}(V(T \otimes T_{\mathcal{A}}))Z(Y)]$, $Y \in \mathcal{F}_{\mathcal{A}}$.

3.2. A measurement scheme $\mathcal{M} := \langle \mathcal{H}_A, Z, T_A, V, f \rangle$ defines an observable $E^{\mathcal{M}}$ of \mathcal{S} with the space of values (Ω, \mathcal{F}) via the relation: for any $X \in \mathcal{F}$ and $T \in \mathcal{S}(\mathcal{H})$,

$$(2) \quad p_T^{E^{\mathcal{M}}}(X) := p_{\mathcal{R}_A(V(T \otimes T_{CalA}))}^Z(f^{-1}(X)).$$

This is the observable *measured* by means of the scheme \mathcal{M} in the sense that the totality of the distributions $p_{\mathcal{R}_A(V(T \otimes T_{CalA}))}^Z$ (for all $T \in \mathcal{S}(\mathcal{H})$) of the pointer outcomes in the final apparatus states determine the POV measure $E^{\mathcal{M}}$ via (2). A measurement scheme \mathcal{M} is a measurement of a given observable E if the measured observable $E^{\mathcal{M}}$ equals E .

3.3. There is an important subclass of measurement schemes for \mathcal{S} . They consist of a sharp pointer observable Z , a vector state preparation of \mathcal{A} , $T_A = P[\phi]$, $\phi \in \mathcal{H}_A$, $\langle \phi | \phi \rangle = 1$, and a unitary measurement coupling $V(T \otimes T_{CalA}) = UT \otimes T_A U^*$, with a unitary map U on $\mathcal{H} \otimes \mathcal{H}_A$. Subsuming the possible pointer function in the definition of Z by identification of Ω and Ω_A , we denote such a scheme $\mathcal{M}_U := \langle \mathcal{H}_A, Z, P[\phi], U \rangle$ and call it a *unitary measurement scheme* (with the understanding that Z is a sharp observable). It is a basic result of the quantum theory of measurement that every observable E of \mathcal{S} admits a unitary measurement, that is, there is a unitary measurement scheme \mathcal{M}_U such that $E^{\mathcal{M}_U} = E$.⁴ Thus the relation between measurement schemes and POV measures induced by (2) defines a map from the former onto the latter. In spite of this fundamental result, it is important to consider general measurement schemes \mathcal{M} , since in many applications the choices of a sharp pointer and a vector state preparation of the apparatus are not physically realizable.

3.4. Another basic condition for a measurement scheme $\langle \mathcal{H}_A, Z, T_A, V, f \rangle$ to qualify as a measurement is the requirement that the measurement should lead to a definite result. We take this requirement to entail, first of all, that the pointer observable Z should have assumed a definite value after the measurement. One should then be able to ‘read the actual value’ of the pointer observable Z and deduce from this the value of the measured observable. As well known, the task of explaining the occurrence of a definite pointer value at the end of a measuring process presents one of the major open problems in quantum mechanics. We do not enter this difficult question here. (For an overview of the issues involved, the reader may wish to consult ref. 1). There are, however, some necessary conditions for the pointer observable Z to assume a definite value in the final apparatus state $\mathcal{R}_A(V(T \otimes T_{CalA}))$, conditions which are tractable and which call for the study of the correlation properties of a measurement. These conditions are the subject of the present paper.

B. Reading of pointer values.

3.5. The reading of measurement outcomes involves the discrimination between the elements of a finite (or, as an idealization, countable) set of alternative pointer values. In order to formulate this idea in the general context of an E -measurement \mathcal{M} , we introduce the notion of a *reading scale* as a countable partition of the value

space of the pointer observable, $\Omega_{\mathcal{A}} = \cup f^{-1}(X_i)$, induced by a countable partition of the value space of the measured observable, $\Omega = \cup X_i$, $X_i \in \mathcal{F}$, $X_i \cap X_j = \emptyset$ for $i \neq j$. Such a reading scale will be denoted \mathcal{R} , and we let \mathbf{I} denote its index set.

3.6. A reading scale \mathcal{R} determines discrete, coarse-grained versions of the pointer observable Z and the measured observable E ;

$$(3a) \quad Z^{\mathcal{R}} : i \mapsto Z_i := Z(f^{-1}(X_i)), \quad i \in \mathbf{I},$$

$$(3b) \quad E^{\mathcal{R}} : i \mapsto E_i := E(X_i), \quad i \in \mathbf{I}.$$

The $Z^{\mathcal{R}}$ -value i refers to the pointer reading $f^{-1}(X_i)$ which, in turn, is correlated to the value set X_i of the measured observable E . If E itself is discrete there is a natural (finest) reading scale \mathcal{R} such that $E = E^{\mathcal{R}}$ and $Z^f = Z^{\mathcal{R}}$. It should be noted that we have included the pointer function f in the definition of $Z^{\mathcal{R}}$ so that the two discrete observables (3) do have the same set of values.

3.7. We say that the (discrete) pointer observable $Z^{\mathcal{R}}$ *has the value i* in the state $\mathcal{R}_{\mathcal{A}}(V(T \otimes T_{CalA}))$ if and only the measurement outcome probability for this value equals one: $\text{tr}[\mathcal{R}_{\mathcal{A}}(V(T \otimes T_{CalA}))Z_i] = 1$. Since $\text{tr}[TE_i] = \text{tr}[\mathcal{R}_{\mathcal{A}}(V(T \otimes T_{CalA}))Z_i]$, and, in general, $0 \neq p_T^E(X_i) \neq 1$, the pointer observable does not have a value at the end of the measurement. It may, however, occur that the state $\mathcal{R}_{\mathcal{A}}(V(T \otimes T_{CalA}))$ is a *mixture of eigenstates of $Z^{\mathcal{R}}$ with the weights $p_T^E(X_i)$* . This is indeed a necessary condition for the assertion that the pointer observable Z has assumed a definite value with respect to a reading scale \mathcal{R} at the end of the measurement \mathcal{M} . We go on to specify this case further.

3.8. We consider a measurement $\langle \mathcal{H}_{\mathcal{A}}, Z, T_{\mathcal{A}}, V, f \rangle$ of E with a fixed reading scale \mathcal{R} . Any X_i , $i \in \mathbf{I}$, defines a (unnormalised) *conditioned state*:

$$(4) \quad V_i(T) := I \otimes Z_i^{1/2} V(T \otimes T_{CalA}) I \otimes Z_i^{1/2},$$

the state of $\mathcal{S} + \mathcal{A}$ on the condition that the pointer observable $Z^{\mathcal{R}}$ has value i . The (trace) norm of this state is $\text{tr}[V_i(T)] = \text{tr}[V(T \otimes T_{CalA})I \otimes Z_i] = p_T^E(X_i)$, and the corresponding (normalised) reduced states, the *final component states* of \mathcal{S} and \mathcal{A} are:

$$(5a) \quad T_{\mathcal{S}}(i, T) := p_T^E(X_i)^{-1} \mathcal{R}_{\mathcal{S}}(V_i(T)),$$

$$(5b) \quad T_{\mathcal{A}}(i, T) := p_T^E(X_i)^{-1} \mathcal{R}_{\mathcal{A}}(V_i(T)).$$

(If $p_T^E(X_i) = 0$, we put $T_{\mathcal{S}}(i, T) = T_{\mathcal{A}}(i, T) = O$). The conditional interpretation of the states (4) and (5) presupposes, however, that the pointer observable $Z^{\mathcal{R}}$ has value i in state $T_{\mathcal{A}}(i, T)$, that is, $\text{tr}[T_{\mathcal{A}}(i, T)Z_i] = 1$ for all i and T whenever $p_T^E(X_i) \neq 0$.⁵ This requirement is always satisfied if the pointer observable is sharp. In general this is a condition to be imposed on the measurement scheme. We call it the *pointer value-definiteness* condition and note that it may be written in either of the following equivalent forms:

$$(6a) \quad \text{tr}[T_{\mathcal{A}}(i, T)Z_i] = 1 \quad (\text{whenever } p_T^E(X_i) \neq 0),$$

$$(6b) \quad Z_i T_{\mathcal{A}}(i, T) = T_{\mathcal{A}}(i, T),$$

for all $i \in \mathbf{I}$ and all initial states T of \mathcal{S} .

3.9. For any reading scale \mathcal{R} and any state $T \in \mathcal{S}(\mathcal{H})$ one has

$$(7) \quad \mathcal{R}_{\mathcal{S}}(V(T \otimes T_{CalA})) = \sum p_T^E(X_i) T_{\mathcal{S}}(i, T);$$

this is to say that the final object state behaves additively with respect to the pointer conditioning: that is, the state of \mathcal{S} on the plain condition that the measurement has been performed, is the same as the state of \mathcal{S} after the measurement conditional on the fact that the pointer value is registered with respect to the reading scale \mathcal{R} .⁶ Although it also holds true that for any $i \in \mathbf{I}$ and $T \in \mathcal{S}(\mathcal{H})$

$$(8) \quad T_{\mathcal{A}}(i, T) = p_T^E(X_i)^{-1} Z_i^{1/2} \mathcal{R}_{\mathcal{A}}(V(T \otimes T_{CalA})) Z_i^{1/2},$$

it is not the case, in general, that the final apparatus state $\mathcal{R}_{\mathcal{A}}(V(T \otimes T_{CalA}))$ is conditioned with respect to \mathcal{R} ; thus, in general,

$$(9) \quad \mathcal{R}_{\mathcal{A}}(V(T \otimes T_{CalA})) \neq \sum p_T^E(X_i) T_{\mathcal{A}}(i, T).$$

The requirement that $\mathcal{R}_{\mathcal{A}}(V(T \otimes T_{CalA}))$ is a mixture of the final component states $T_{\mathcal{A}}(i, T)$ is therefore another condition on the measurement²; we call it the *pointer mixture condition*:

$$(10) \quad \mathcal{R}_{\mathcal{A}}(V(T \otimes T_{\mathcal{A}})) = \sum p_T^E(X_i) T_{\mathcal{A}}(i, T)$$

for all initial states T of \mathcal{S} .

3.10. The pointer value-definiteness condition (6) and the pointer mixture condition (10) imply that the final apparatus state is a mixture of the pointer eigenstates $T_{\mathcal{A}}(i, T)$ with the weights $p_T^E(X_i)$; this means that the final apparatus state $\mathcal{R}_{\mathcal{A}}(V(T \otimes T_{CalA}))$ is conditioned with respect to the reading scale \mathcal{R} .^{2,6} One may consider the assumption that in addition to this, the state $\mathcal{R}_{\mathcal{A}}(V(T \otimes T_{CalA}))$ admits the *ignorance interpretation* with respect to the decomposition (10): that is, the apparatus [in state $\mathcal{R}_{\mathcal{A}}(V(T \otimes T_{CalA}))$] is actually in one of the component states $T_{\mathcal{A}}(i, T)$, and this is the case with the subjective probability $p_T^E(X_i)$. As well known, such an interpretation of the mixed state $\mathcal{R}_{\mathcal{A}}(V(T \otimes T_{CalA}))$ is extremely problematic and in most cases impossible; but if it were the case then the pointer could be claimed to have a definite value i (with respect to a reading scale \mathcal{R}) after the measurement with the subjective probability $p_T^E(X_i)$. Setting aside the difficulties with the ignorance interpretation (and thus with explaining the occurrence of definite measurement outcomes in quantum mechanics), it still is important to investigate more closely the conditions (6) and (10) and to see how these possible properties of a measurement are related to the structure of the final state of the object system.

3.11 THEOREM. *Let \mathcal{M} be a measurement of an observable E and \mathcal{R} any reading scale. For any initial state T of the object system, the condition a) implies the conditions b) and c):*

- a) $T_{\mathcal{S}}(i, T) \cdot T_{\mathcal{S}}(j, T) = O$ for $i \neq j$;
- b) $\mathcal{R}_{\mathcal{A}}(V(T \otimes T_{\mathcal{A}})) = \sum p_T^E(X_i) T_{\mathcal{A}}(i, T)$ for all i ;
- c) $Z_i T_{\mathcal{A}}(i, T) = T_{\mathcal{A}}(i, T)$ for all i .

If \mathcal{M} is a unitary measurement \mathcal{M}_U , then a) and b) are equivalent conditions for any initial vector state $T = P[\varphi]$ of \mathcal{S} .

PROOF:

a) \Rightarrow b)&c): For each i , let F_i be the support projection of $T_{\mathcal{S}}(i, T)$, that is, the smallest projection Q such that $QT_{\mathcal{S}}(i, T) = T_{\mathcal{S}}(i, T)$. Then one gets (for $i \neq j$):

$$\begin{aligned}
 T_{\mathcal{S}}(i, T) \cdot T_{\mathcal{S}}(j, T) = O &\Leftrightarrow F_i T_{\mathcal{S}}(j, T) = O \\
 &\Leftrightarrow \text{tr}[F_i \otimes I V_j(T)] = 0 \\
 &\Leftrightarrow F_i \otimes Z_j^{1/2} V(T \otimes T_{\mathcal{A}}) I \otimes Z_j^{1/2} = O \\
 &\Leftrightarrow F_i \otimes Z_j^{1/2} V(T \otimes T_{\mathcal{A}})^{1/2} = O \\
 (\alpha) \quad &\Rightarrow F_i \otimes Z_j V(T \otimes T_{\mathcal{A}})^{1/2} = O
 \end{aligned}$$

By the definition of F_i one also has

$$\begin{aligned}
 F_i T_{\mathcal{S}}(i, T) &= T_{\mathcal{S}}(i, T) \\
 &\Leftrightarrow F_i \otimes Z_i^{1/2} V(T \otimes T_{\mathcal{A}}) I \otimes Z_i^{1/2} = I \otimes Z_i^{1/2} V(T \otimes T_{\mathcal{A}}) I \otimes Z_i^{1/2} \\
 (\beta) \quad &\Rightarrow F_i \otimes Z_i V(T \otimes T_{\mathcal{A}})^{1/2} = I \otimes Z_i V(T \otimes T_{\mathcal{A}})^{1/2}
 \end{aligned}$$

Combining (α) and (β) and using the fact that $\sum Z_i = I$ yields

$$(\gamma) \quad F_i \otimes I V(T \otimes T_{\mathcal{A}})^{1/2} = F_i \otimes Z_i V(T \otimes T_{\mathcal{A}})^{1/2} = I \otimes Z_i V(T \otimes T_{\mathcal{A}})^{1/2}.$$

From this one obtains $I \otimes Z_i V(T \otimes T_{\mathcal{A}}) = I \otimes Z_i^2 V(T \otimes T_{\mathcal{A}})$, which gives c). Using c), one shows similarly that

$$I \otimes Z_i V(T \otimes T_{\mathcal{A}})^{1/2} = I \otimes Z_i^{1/2} V(T \otimes T_{\mathcal{A}})^{1/2}.$$

Inserting this in (γ) , multiplying each term with its adjoint and summing over i , one obtains

$$\sum F_i \otimes I V(T \otimes T_{\mathcal{A}}) F_i \otimes I = \sum I \otimes Z_i^{1/2} V(T \otimes T_{\mathcal{A}}) I \otimes Z_i^{1/2}.$$

Taking the partial trace with respect to \mathcal{H} finally yields b).

b) \Rightarrow a): This implication will be shown for a unitary measurement \mathcal{M}_U and for

vector state preparations $T = P[\varphi]$. In that case $T_{\mathcal{A}} = P[\phi]$ and $V(T \otimes T_{\mathcal{A}}) = P[U(\varphi \otimes \phi)]$. Denoting the biorthogonal decomposition of this state as $U(\varphi \otimes \phi) = \sum_{nk} c_n \varphi_{nk} \otimes \phi_{nk}$, with $c_n > 0$, we obtain $T_{\mathcal{A}}(\Omega, T) = \sum_{nk} |c_n|^2 P[\phi_{nk}]$. Now $b)$ implies that all the projections Z_i commute with $T_{\mathcal{A}}(\Omega, T)$. Therefore one can choose the orthonormal system $\{\phi_{nk}\}$ such that $Z_i \phi_{nk} = \phi_{nk}$ or $Z_i \phi_{nk} = 0$. Thus there is a renumbering of this system, $\{\phi_{nk}\} = \{\phi_{il}\}$, such that $Z_i \phi_{il} = \phi_{il}$. It follows that there are corresponding renumberings $\{\varphi_{il}\} = \{\varphi_{nk}\}$ and $\{d_{il}\} = \{c_n\}$ such that $U(\varphi \otimes \phi) = \sum d_{il} \varphi_{il} \otimes \phi_{il}$. Then $T_{\mathcal{S}}(i, T) = \sum_{\ell} |d_{i\ell}|^2 P[\varphi_{i\ell}]$. Since the subsets of vectors $\varphi_{i\ell}$ with different values of i are mutually disjoint and therefore orthogonal, one concludes that $a)$ holds. This completes the proof.

It can be demonstrated by means of examples that the implication $b) \Rightarrow a)$ need not hold if the measurement is not unitary or if the initial pointer state is not pure.¹

C. First kind and repeatable measurements.

3.12. A measurement \mathcal{M} of an observable E is of the *first kind* if the probability for a given result is the same both before and after the measurement, that is, for any $T \in \mathcal{S}(\mathcal{H})$ and for all $X \in \mathcal{F}$,

$$(11) \quad p_T^E(X) = p_{\mathcal{R}_{\mathcal{S}}(V(T \otimes T_{\text{CalA}}))}^E(X).$$

Unitary measurement schemes with a coupling $U = e^{i\lambda A \otimes B}$, $\lambda \in \mathbf{R}$, A (on \mathcal{H}) and B (on $\mathcal{H}_{\mathcal{A}}$) self-adjoint, do give rise to such measurements; we refer to Sec. 8 for an analysis of this model.

3.13. A measurement \mathcal{M} of an observable E is *repeatable* if its repetition does not lead to a new result. One way to express the requirement is the following: for any $T \in \mathcal{S}(\mathcal{H})$ and $X \in \mathcal{F}$, if $p_T^E(X) \neq 0$, then

$$(12) \quad p_{T_{\mathcal{S}}(X, T)}^E(X) = 1,$$

(where $T_{\mathcal{S}}(X, T)$ is defined by (3a), (4) and (5a) with $X = X_i$). Equivalently, \mathcal{M} is a repeatable E -measurement if for any $T \in \mathcal{S}(\mathcal{H})$ and $X \in \mathcal{F}$, for which $p_T^E(X) \neq 0$, it holds true that

$$(13) \quad E(X) T_{\mathcal{S}}(X, T) = T_{\mathcal{S}}(X, T).$$

Another basic result of measurement theory is that an observable E which admits a repeatable measurement is discrete.^{4,7}

3.14. According to (13), a repeatable measurement drives the object system into an eigenstate of the measured observable $E : i \mapsto E_i$. The orthogonality conditions of Theorem 3.11 are then satisfied and the final apparatus state $\mathcal{R}_{\mathcal{A}}(V(T \otimes T_{\text{CalA}}))$ is the mixture of the eigenstates $T_{\mathcal{A}}(i, T)$ of $Z : i \mapsto Z_i$ with the weights $\text{tr}[TE_i]$.

3.15. It is evident that repeatable measurements are also of the first kind. However, as will be demonstrated in Sec. 8, a first kind measurement need not be repeatable, though for sharp observables the two notions coincide.⁸

IV. Statistical dependence and correlations. A measurement \mathcal{M} of an observable E brings the compound object-apparatus system into an entangled state $V(T \otimes T_{\mathcal{A}})$. The possibility of transferring information from \mathcal{A} to \mathcal{S} rests on the fact that this state entails *statistical dependencies* between quantities pertaining to these systems. Accordingly, three types of *correlations* inherent in the state $V(T \otimes T_{\mathcal{A}})$ are of special interest for characterising the measurement: *i*) correlations between the measured observable and the pointer observable; *ii*) correlations between the corresponding values of these observables; and *iii*) correlations between the final component states of the two subsystems. For their study it is helpful to recall some basic notions and facts concerning the relation between statistical dependence and correlation.

4.1. Let μ be a probability measure on the real Borel space $(\mathbf{R}^2, \mathcal{B}(\mathbf{R}^2))$, and let μ_1 and μ_2 be the marginal measures of μ with respect to a Cartesian coordinate system: for $X, Y \in \mathcal{B}(\mathbf{R})$,

$$(14) \quad \mu_1(X) = \mu(X \times \mathbf{R}), \quad \mu_2(Y) = \mu(\mathbf{R} \times Y).$$

These marginal measures correspond to the coordinate projections (random variables) $\pi_1 : (x, y) \mapsto x$ and $\pi_2 : (x, y) \mapsto y$ in the sense that $\mu_i = \mu^{\pi_i}$, that is, $\mu_i(X) = \mu^{\pi_i}(X) := \mu(\pi_i^{-1}(X))$ for all $X \in \mathcal{B}(\mathbf{R})$, $i = 1, 2$. Assume that the expectations and the variances of μ_i are well defined and finite: $\epsilon_i = \int x d\mu_i(x)$, $\sigma_i^2 = \int (x - \epsilon_i)^2 d\mu_i(x)$, and let $\epsilon_{12} = \int xy d\mu(x, y)$. The (normalised) *correlation* of the marginal measures μ_1 and μ_2 in μ is then defined as:

$$(15) \quad \rho(\mu_1, \mu_2; \mu) := \int \frac{(x - \epsilon_1)(y - \epsilon_2)}{\sigma_1 \sigma_2} d\mu(x, y) = \frac{\epsilon_{12} - \epsilon_1 \epsilon_2}{\sigma_1 \sigma_2}$$

(whenever $\sigma_1 \neq 0 \neq \sigma_2$). The Schwarz inequality entails $|\rho(\mu_1, \mu_2; \mu)| \leq 1$. The marginals μ_1, μ_2 are *uncorrelated* if $\rho(\mu_1, \mu_2; \mu) = 0$ (that is, $\epsilon_{12} = \epsilon_1 \epsilon_2$), *strongly correlated* if $\rho(\mu_1, \mu_2; \mu) = 1$ (that is, $\epsilon_{12} - \epsilon_1 \epsilon_2 = \sigma_1 \sigma_2$), and *strongly anticorrelated* if $\rho(\mu_1, \mu_2; \mu) = -1$ (that is, $\epsilon_{12} - \epsilon_1 \epsilon_2 = -\sigma_1 \sigma_2$). The strong correlation conditions can also be written in terms of the coordinate projections π_1 and π_2 :

$$(16a) \quad \rho(\pi_1, \pi_2; \mu) = +1 \quad \text{iff} \quad \pi_1 = \frac{\sigma_1}{\sigma_2}(\pi_2 - \epsilon_2) + \epsilon_1 =: \ell_+ \circ \pi_2 \quad (\mu\text{-a.e.}),$$

$$(16b) \quad \rho(\pi_1, \pi_2; \mu) = -1 \quad \text{iff} \quad \pi_1 = -\frac{\sigma_1}{\sigma_2}(\pi_2 - \epsilon_2) + \epsilon_1 =: \ell_- \circ \pi_2 \quad (\mu\text{-a.e.}).$$

(Here we have introduced the function $\ell_{\pm} : y \mapsto \ell_{\pm}(y) := \pm \frac{\sigma_1}{\sigma_2}(y - \epsilon_2) + \epsilon_1$). A case of special interest arises when the marginals μ_1 and μ_2 have the same (finite) first and second moments so that $\epsilon_1 = \epsilon_2$, $\sigma_1 = \sigma_2$. Then one has:

$$(17a) \quad \begin{aligned} \rho(\mu_1, \mu_2; \mu) = +1 \quad &\text{iff} \quad \epsilon_{12} = \epsilon_1^2 + \sigma_1^2, \\ &\text{iff} \quad \pi_1 = \pi_2 \quad (\mu\text{-a.e.}), \end{aligned}$$

$$(17b) \quad \begin{aligned} \rho(\mu_1, \mu_2; \mu) = -1 \quad &\text{iff} \quad \epsilon_{12} = \epsilon_1^2 - \sigma_1^2, \\ &\text{iff} \quad \pi_1 = -\pi_2 + 2\epsilon_1 \quad (\mu\text{-a.e.}). \end{aligned}$$

4.2. The notion of correlation can be applied to quantify the degree of mutual dependence of the marginal measures. In order to avoid dealing with unnecessary complications, we assume that μ_1 and μ_2 are no $\{0,1\}$ -valued measures; equivalently, we let $\sigma_1 \neq 0 \neq \sigma_2$. μ_1 and μ_2 are *independent* if $\mu = \mu_1 \times \mu_2$. Otherwise, μ_1, μ_2 are *dependent*. They are *completely dependent* if there is a (measurable) function $h : \mathbf{R} \rightarrow \mathbf{R}$ such that $\mu(X \times Y) = \mu_2(h^{-1}(X) \cap Y)$ for $X, Y \in \mathcal{B}(\mathbf{R})$. That is, the marginal measure μ_2 suffices to determine the whole measure μ . The relation of complete dependence is symmetric with respect to the two marginals only if h is bijective. This is the case of concern here.

4.3. It is evident that the statistical independence of μ_1, μ_2 implies $\rho(\mu_1, \mu_2; \mu) = 0$. However, the latter condition is not sufficient to ensure their independence. (For a counter example, see, for instance ref. 9). On the other hand, eqs. (16a,b) show that strong (anti)correlation entails complete dependence, the dependence being given by the linear function ℓ_{\pm} . Indeed, the condition $\pi_1 = \ell_{\pm} \circ \pi_2$ (μ -a.e.) implies that $\mu(X \times Y) = 0$ for all X and Y for which $\ell_{\pm}^{-1}(X) \cap Y = \emptyset$. Thus, in particular, for any X and Y , and with X' denoting the complement of X one has $\mu(X' \times \ell_{\pm}^{-1}(X) \cap Y) = 0 = \mu(X \times \ell_{\pm}^{-1}(X') \cap Y)$. The additivity properties of μ allow one then to verify that for all X, Y , $\mu_2(\ell_{\pm}^{-1}(X) \cap Y) = \mu(X \times Y)$, that is, μ_1 and μ_2 are completely dependent with ℓ_{\pm} . By a direct computation one can confirm that the converse implication holds true whenever the function h is linear. Therefore, we have:

$$(18a) \quad \begin{aligned} \rho(\mu_1, \mu_2; \mu) &= +1 \quad \text{iff } \mu_1, \mu_2 \text{ are completely dependent} \\ &\quad \text{with } h(y) = ay + b, \quad a > 0, \end{aligned}$$

$$(18b) \quad \begin{aligned} \rho(\mu_1, \mu_2; \mu) &= -1 \quad \text{iff } \mu_1, \mu_2 \text{ are completely dependent} \\ &\quad \text{with } h(y) = ay + b, \quad a < 0. \end{aligned}$$

In both cases the constants are $a = \pm\sigma_1/\sigma_2$, $b = \epsilon_1 - a\epsilon_2$, so that $h = \ell_{\pm}$.

V. Strong correlations between observables.

5.1. According to the condition (2), in an E -measurement the initial E -outcome distribution is recovered from the final Z -outcome distribution. In addition to this basic requirement, a measurement may also establish complete statistical dependence between the measured observable and the pointer observable after the measurement; that is, the observables E and Z^f may become strongly correlated in the final object-apparatus state $V(T \otimes T_A)$. In order to avoid technical complications in the formulation of this correlation, we assume that the value space of E is the real Borel space, $(\Omega, \mathcal{F}) = (\mathbf{R}, \mathcal{B}(\mathbf{R}))$. Then for any state $T \in \mathcal{S}(\mathcal{H})$ the map

$$(19) \quad \mu : X \times Y \mapsto \text{tr}[V(T \otimes T_A)E(X) \otimes Z^f(Y)]$$

extends to a probability measure on $(\mathbf{R}^2, \mathcal{B}(\mathbf{R}^2))$.¹⁰ The marginal distributions are

$$(20a) \quad \mu_1 : X \mapsto \text{tr}[\mathcal{R}_{\mathcal{S}}(V(T \otimes T_{CalA}))E(X)],$$

$$(20b) \quad \mu_2 : Y \mapsto \text{tr}[\mathcal{R}_{\mathcal{A}}(V(T \otimes T_{CalA}))Z^f(Y)] = \text{tr}[TE(Y)].$$

Denoting the correlation of μ_1 and μ_2 in μ as $\rho(E, Z^f; V(T \otimes T_{\mathcal{A}}))$, we say that the measurement \mathcal{M} of E produces *strong observable-(anti)correlation* in state T if this number equals 1 (−1). According to (18), this occurs exactly when the probability measures (20a,b) are completely dependent, with the function ℓ_{\pm} . In order to analyze the statistical dependence of μ_1 and μ_2 we shall make use of the concept of a state transformer (also known as an instrument) associated with a measurement.

5.2. Consider a measurement $\langle \mathcal{H}_{\mathcal{A}}, Z, T_{\mathcal{A}}, V, f \rangle$ of E . Any $X \in \mathcal{F}$ defines a non-normalised state

$$(21) \quad V_X(T) := I \otimes Z^{1/2} (f^{-1}(X)) V(T \otimes T_{\text{CalA}}) I \otimes Z^{1/2} (f^{-1}(X)),$$

the (trace) norm of which is $\text{tr}[V_X(T)] = p_T^E(X)$. Taking the partial trace of $V_X(T)$ over $\mathcal{H}_{\mathcal{A}}$ one gets the (nonnormalised) reduced state of \mathcal{S} ,

$$(22) \quad \mathcal{I}_X(T) := \mathcal{R}_{\mathcal{S}}(V_X(T)).$$

For any $X \in \mathcal{F}$ and $T \in \mathcal{S}(\mathcal{H})$, $\text{tr}[\mathcal{I}_X(T)] = \text{tr}[TE(X)]$, and $T \mapsto \mathcal{I}_X(T)$ is a (contractive) state transformation. The mapping $\mathcal{I} : X \mapsto \mathcal{I}_X$ has the measure property $\text{tr}[\mathcal{I}_{\cup X_i}(T)] = \sum \text{tr}[\mathcal{I}_{X_i}(T)]$ for any disjoint sequence $(X_i) \subset \mathcal{F}$ and for all $T \in \mathcal{S}(\mathcal{H})$. Moreover, $\text{tr}[\mathcal{I}_{\Omega}(T)] = 1$ for any T . We call \mathcal{I} the *state transformer* induced by the measurement \mathcal{M} . It describes the object system's state changes under the measurement, and it uniquely defines the measured observable via the relation $\text{tr}[\mathcal{I}_X(T)] = \text{tr}[TE(X)]$. We note also that $p_T^E(X) T_{\mathcal{S}}(X, T) = \mathcal{I}_X(T)$, and, in particular, $\mathcal{R}_{\mathcal{S}}(V(T \otimes T_{\text{CalA}})) = \mathcal{I}_{\Omega}(T)$.

5.3. The probability measure (19) can be written as

$$(23) \quad \mu(X \times Y) = \text{tr}[\mathcal{I}_Y(T)E(X)] = \text{tr}[\mathcal{I}_X(\mathcal{I}_Y(T))],$$

and the second marginal is $\mu_2(Y) = \text{tr}[\mathcal{I}_Y(T)]$. The strong (anti-)correlation then amounts to

$$(24) \quad \text{tr}[\mathcal{I}_X(\mathcal{I}_Y(T))] = \text{tr}[\mathcal{I}_{\ell_{\pm}^{-1}(X) \cap Y}(T)].$$

A special case of complete dependence arises with ℓ_+ being the identity function:

$$(25) \quad \text{tr}[\mathcal{I}_X(\mathcal{I}_Y(T))] = \text{tr}[\mathcal{I}_{X \cap Y}(T)].$$

This relation is easily seen to coincide with (12).⁸ Thus, if valid for all states T , (25) expresses the repeatability of the measurement, and we may conclude that any repeatable measurement leads to strong observable-correlations. The repeatability condition (25) is not necessary for the strong observable-correlation (24).

5.4. Condition (25) implies, in particular, the equality of the marginal measures μ_1, μ_2 of Eqs. (20a,b): for all X ,

$$(26) \quad p_T^E(X) = p_{\mathcal{R}_{\mathcal{S}}(V(T \otimes T_{\mathcal{A}}))}^E(X).$$

This is just the first-kind property of the measurement. It may occur that these marginal measures coincide irrespectively of whether (25) holds or not; in that case conditions (17a,b) give the relevant characterisations of strong (anti)correlations.

5.5 THEOREM. *Let \mathcal{M} be a measurement of an observable E , and let \mathcal{R} be any reading scale. Then a) implies b), where:*

- a) $E(X_i)T_{\mathcal{S}}(i, T) = T_{\mathcal{S}}(i, T)$ for all $T \in \mathcal{S}(\mathcal{H})$, $X_i \in \mathcal{R}$;
- b) $\sigma(p_{\mathcal{R}_{\mathcal{S}}(V(T \otimes T_{\mathcal{A}}))}^{E^{\mathcal{R}}}) \neq 0$ and $\rho(E^{\mathcal{R}}, Z^{\mathcal{R}}; V(T \otimes T_{\mathcal{A}})) = 1$
for all $T \in \mathcal{S}(\mathcal{H})$ with $\sigma(p_T^{E^{\mathcal{R}}}) \neq 0$.

If the reading scale \mathcal{R} is finite, then a) and b) are equivalent.

PROOF: The eigenstate condition a) is equivalent with the repeatability condition (with respect to \mathcal{R}). Therefore, if a) holds, then also b) is true. It remains to show that b) implies a) whenever \mathcal{R} is finite. According to (18a), the statement $\rho(E^{\mathcal{R}}, Z^{\mathcal{R}}; V(T \otimes T_{\mathcal{A}})) = 1$ is equivalent to the complete dependence, $\mu(i, j) = \mu_2(j) \delta_{i, \ell_+(j)}$, with a bijective linear mapping $i = \ell_+(j) = aj + b$, $a > 0$, between those values i, j for which $\mu_2(j) \neq 0$ (and hence $\mu_1(i) = \mu_2(\frac{1}{a}(i - b)) \neq 0$).

Case 1. Let T be such that $0 \neq \text{tr}[TE_i] \neq 1$ for all $i \in \mathbf{I}$. Then $\mu(i, j)$ correlates, via $i = \ell_+(j) = aj + b$, all values $j \in \mathbf{I}$ with values $i \in \mathbf{I}$. Since ℓ_+ is onto and monotonically increasing, $\ell_+(j) = j$. But the complete dependence condition, with $\ell_+(j) = j$, is nothing but Eq. (25) (with respect to \mathcal{R}), which is equivalent to a).

Case 2. Let T be any state such that $0 \neq \text{tr}[TE_k] \neq 1$ holds exactly for all $k \in \mathbf{I}_1$, a proper nonempty subset of \mathbf{I} . Take any T' for which $0 \neq \text{tr}[T'E_l] \neq 1$ exactly for all $l \in \mathbf{I}'_1$, the complement of \mathbf{I}_1 . Then the reasoning of Case 1 applies to $\hat{T} := \frac{1}{2}T + \frac{1}{2}T'$. Hence, $E_i \mathcal{I}_i(\hat{T}) = \mathcal{I}_i(\hat{T})$ for all $i \in \mathbf{I}$. Inserting in this equation the relation $\mathcal{I}_k(T') = O$, which holds for $k \in \mathbf{I}_1$, it follows that $E_i \mathcal{I}_i(T) = \mathcal{I}_i(T)$ for $i \in \mathbf{I}_1$. But this relation holds trivially also for $i \in \mathbf{I}'_1$ since in that case $\mathcal{I}_i(T) = O$. This completes the proof.

VI. Strong correlations between values.

6.1. The observable $E^{\mathcal{R}}$ measured by the scheme \mathcal{M} with the reading scale \mathcal{R} is discrete. One may therefore ask to what degree the *values* of this observable and the pointer observable $Z^{\mathcal{R}}$ become correlated in the measurement. To answer this question requires studying the correlation $\rho(E_i, Z_i; V(T \otimes T_{\mathcal{A}}))$ of the i -th values of these observables in the final object-apparatus state, that is, the correlation of quantities $E_i \otimes I$ and $I \otimes Z_i$ in the state $V(T \otimes T_{\mathcal{A}})$:

$$(27) \quad \rho(E_i, Z_i; V(T \otimes T_{\mathcal{A}})) = \frac{\epsilon_{12} - \epsilon_1 \epsilon_2}{\sigma_1 \sigma_2}.$$

The respective quantities are easily determined:

$$(28a) \quad \epsilon_{12} = \text{tr}[\mathcal{I}_i^2(T)],$$

$$(28b) \quad \epsilon_1 = \text{tr}[\mathcal{I}(\mathbf{I})(T)E_i],$$

$$(28c) \quad \epsilon_2 = \text{tr}[TE_i],$$

$$(28d) \quad \sigma_1^2 = \text{tr}[\mathcal{I}(\mathbf{I})(T)E_i^2] - \text{tr}[\mathcal{I}(\mathbf{I})(T)E_i]^2,$$

$$(28e) \quad \sigma_2^2 = \text{tr}[\mathcal{R}_{\mathcal{A}}(V(T \otimes T_{\mathcal{A}}))Z_i^2] - \text{tr}[TE_i]^2.$$

Strong correlation is then equivalent to

$$(29) \quad \epsilon_{12} - \epsilon_1 \epsilon_2 = \sigma_1 \sigma_2$$

whenever the right-hand side is nonzero.

6.2. Assume that the final component state $T_S(i, T)$ is a 1-eigenstate of E_i (whenever $p_T^E(X_i) \neq 0$); then one obtains $\epsilon_{12} = \epsilon_1 = \epsilon_2$ for all T . It follows that $\epsilon_{12} - \epsilon_1 \epsilon_2 = \sigma_1^2 \leq \sigma_1 \sigma_2$ and thus $\sigma_1 \leq \sigma_2$. On the other hand, the relation $\epsilon_1 = \epsilon_2 = \epsilon_{12}$ together with $\sigma_2^2 \leq \epsilon_2 - \epsilon_2^2 = \epsilon_{12} - \epsilon_1 \epsilon_2 = \sigma_1^2$ implies $\sigma_2 \leq \sigma_1$. Therefore the correlation $\rho(E_i, Z_i; V(T \otimes T_A))$ equals 1 whenever $0 \neq p_T^E(X_i) \neq 1$.

Another interesting implication of the eigenstate condition $\epsilon_{12} = \epsilon_2$ and the ensuing equality $\sigma_2 = \epsilon_2 - \epsilon_2^2$ is the fact that the state $T_A(i, T)$ is a 1-eigenstate of Z_i . With these observations we have established the following result.

6.3 THEOREM. *Let \mathcal{M} be a measurement of an observable E and let \mathcal{R} be any reading scale. Then for any state T of \mathcal{S} , a) implies b) and c):*

- a) $E_i T_S(i, T) = T_S(i, T)$ for each i ;
- b) $\sigma(E_i \otimes I; V(T \otimes T_A)) \neq 0$ and $\rho(E_i, Z_i; V(T \otimes T_A)) = 1$
for each i with $0 \neq p_T^E(X_i) \neq 1$;
- c) $T_A(i, T)$ is a 1-eigenstate of Z_i for each i with $p_T^E(X_i) \neq 0$.

This result entails that a repeatable measurement is a strong value-correlation measurement. Moreover, a necessary condition for \mathcal{M} to be a repeatable measurement is that the final component state $T_A(i, T)$ of \mathcal{A} is a 1-eigenstate of the pointer observable, that is, \mathcal{M} must fulfil the pointer value-definiteness condition. We recall that this last property and in addition the pointer mixture property arise already as consequences of the mutual orthogonality of the component states $T_S(i, T)$ of \mathcal{S} (Theorem 3.11). The notion of a correlation between values suggests that the observables in question do have definite values; yet it turns out that strong value-correlation does not require pointer value-definiteness, nor repeatability. Even the combination of b) and c) does not require the property a) to hold, as can be demonstrated by simple examples.¹

6.4 THEOREM. *Let \mathcal{M} be a measurement of a sharp observable E and \mathcal{R} any reading scale. For any initial state T of \mathcal{S} , a) is equivalent to b)&c):*

- a) $E_i T_S(i, T) = T_S(i, T)$ for each i ;
- b) $\sigma(E_i \otimes I; V(T \otimes T_A)) \neq 0$ and $\rho(E_i, Z_i; V(T \otimes T_A)) = 1$
for each i with $0 \neq p_T^E(X_i) \neq 1$;
- c) $T_A(i, T)$ is a 1-eigenstate of Z_i for each i with $p_T^E(X_i) \neq 0$.

PROOF: In view of Theorem 6.3 we only need to show that b)&c) implies a). Hence let $\epsilon_{12} - \epsilon_1 \epsilon_2 = \sigma_1 \sigma_2$ hold for each i . Condition c) implies $\sigma_2^2 = \epsilon_2 - \epsilon_2^2$. Similarly

the relation $E_i^2 = E_i$ implies $\sigma_1^2 = \epsilon_1 - \epsilon_1^2$. From Eqs. (28) we obtain $\epsilon_{12} \leq \epsilon_1$, $\epsilon_{12} \leq \epsilon_2$, and therefore

$$\sigma_1\sigma_2 = \epsilon_{12} - \epsilon_1\epsilon_2 \leq \sigma_1^2, \quad \sigma_1\sigma_2 = \epsilon_{12} - \epsilon_1\epsilon_2 \leq \sigma_2^2.$$

This implies $\sigma_1 = \sigma_2$. On the other hand,

$$\epsilon_1\epsilon_2 + \sigma_1\sigma_2 = \epsilon_{12} \leq \epsilon_1 = \sigma_1^2 + \epsilon_1^2, \quad \epsilon_1\epsilon_2 + \sigma_1\sigma_2 = \epsilon_{12} \leq \epsilon_2 = \sigma_2^2 + \epsilon_2^2.$$

Using $\sigma_1 = \sigma_2$, one concludes that $\epsilon_1 = \epsilon_2 = \epsilon_{12}$. But the last equation is equivalent to a). This completes the proof.

VII. Strong correlations between final component states.

7.1. In the two preceding sections it was demonstrated in which way strong observable and value correlations serve as characterisations of repeatable measurements. The corresponding eigenstate condition $E_i T_S(i, T) = T_S(i, T)$ entails, in particular, that the final component states of the object associated with different outcomes i, j are mutually orthogonal, $T_S(i, T) \cdot T_S(j, T) = 0$. In some cases this orthogonality can be characterised in terms of strong correlations between the final component states of \mathcal{S} and \mathcal{A} .

Consider a measurement scheme \mathcal{M} of an observable E with respect to a reading scale \mathcal{R} . We say that \mathcal{M} , with \mathcal{R} , is a *strong state-(anti)correlation* measurement of E if for each initial state T of \mathcal{S} it correlates strongly the final component states $T_S(i, T)$ and $T_A(i, T)$ of the object and the apparatus. This calls for the study of the correlation $\rho(T_S(i, T), T_A(i, T); V(T \otimes T_A))$ of the probability measure defined by the self-adjoint operators $T_S(i, T) \otimes I$ and $I \otimes T_A(i, T)$ and the final object-apparatus state $V(T \otimes T_A)$.

7.2 THEOREM. *Let \mathcal{M} be a measurement of an observable E and \mathcal{R} any reading scale. For any initial state T of the object system for which the component states $T_S(i, T)$ and $T_A(i, T)$ are vector states, a) is equivalent to b)&c):*

- a) $T_S(i, T) \cdot T_S(j, T) = 0$ for $i \neq j$;
- b) $\rho(T_S(i, T), T_A(i, T); V(T \otimes T_A)) = 1$ for each i with $0 \neq p_T^E(X_i) \neq 1$;
- c) $T_A(i, T)$ is a 1-eigenstate of Z_i for each i with $0 \neq p_T^E(X_i) \neq 1$.

PROOF: The equivalence is shown to hold under the assumptions $T_S(i, T) = P[\varphi_i]$ and $T_A(i, T) = P[\phi_i]$. These two relations imply that $I \otimes Z_i^{1/2} V(T \otimes T_A) I \otimes Z_i^{1/2}$ is a vector state of the product form, that is,

$$(\alpha) \quad I \otimes Z_i^{1/2} V(T \otimes T_A) I \otimes Z_i^{1/2} = p_T^E(X_i) P[\varphi_i \otimes \phi_i].$$

If a) holds, then by Theorem 3.11, \mathcal{M} fulfils the pointer value-definiteness condition c). Thus for both implications one can make use of the fact that $Z_i \phi_i = \phi_i$. Then (α) implies

$$I \otimes P[\phi_i] V(T \otimes T_A) I \otimes P[\phi_i] = p_T^E(X_i) P[\varphi_i \otimes \phi_i].$$

With this one computes:

$$\begin{aligned}\epsilon_{12} &= \text{tr}[P[\varphi_i] \otimes P[\phi_i] V(T \otimes T_{\mathcal{A}})] = p_T^E(X_i), \\ \epsilon_1 &= \text{tr}[P[\varphi_i] \mathcal{R}_{\mathcal{S}}(V(T \otimes T_{\mathcal{A}}))] = \sum_j p_T^E(X_j) \text{tr}[P[\varphi_i] P[\varphi_j]], \\ \epsilon_2 &= \text{tr}[I \otimes P[\phi_i] V(T \otimes T_{\mathcal{A}})] = p_T^E(X_i), \\ \sigma_1^2 &= \epsilon_1 - \epsilon_1^2, \\ \sigma_2^2 &= \epsilon_2 - \epsilon_2^2.\end{aligned}$$

$a) \Rightarrow b)$: $a)$ is equivalent to $\text{tr}[P[\varphi_i] P[\varphi_j]] = \delta_{ij}$, one has $\epsilon_1 = \epsilon_2 = \epsilon_{12}$, and $\sigma_1 = \sigma_2$. Thus $\epsilon_{12} - \epsilon_1 \epsilon_2 = \sigma_1 \sigma_2$, that is, $b)$.

$b) \& c) \Rightarrow a)$: Let $\epsilon_{12} - \epsilon_1 \epsilon_2 = \sigma_1 \sigma_2$. Using the inequalities $\epsilon_{12} - \epsilon_1 \epsilon_2 \leq \sigma_k^2$, $k = 1, 2$, one concludes that $\sigma_1 = \sigma_2$. Since $\epsilon_{12} = \epsilon_2$, one also has $\epsilon_{12} - \epsilon_1 \epsilon_2 = \sigma_2^2 = \epsilon_2 - \epsilon_2^2$, and therefore $\epsilon_1 = \epsilon_2$. But from the definition of ϵ_1 one has $\epsilon_1 \geq \epsilon_2$, so that the equality of these numbers implies $\text{tr}[P[\varphi_i] P[\varphi_j]] = 0$ whenever $i \neq j$, that is $a)$.

This completes the proof.

7.3. One may also ask whether the requirement of strong correlation between the final \mathcal{S} and \mathcal{A} states $\mathcal{R}_{\mathcal{S}}(V(T \otimes T_{\mathcal{A}}))$ and $\mathcal{R}_{\mathcal{A}}(V(T \otimes T_{\mathcal{A}}))$ imposes any constraint on the measurement scheme under consideration. That this cannot be expected in general can be seen in the case of a unitary measurement \mathcal{M}_U . Note first that the reduced states of $P[U(\varphi \otimes \phi)]$ have the same spectra, including multiplicities. The spectral decompositions can be given in terms of orthonormal systems $\{\varphi_i\}$, $\{\phi_i\}$ defined by the biorthogonal decomposition $U(\varphi \otimes \phi) = \sum_i c_i \varphi_i \otimes \phi_i$ ($c_i > 0$), and a straightforward calculation shows that

$$(30) \quad \rho(\mathcal{R}_{\mathcal{S}}(P[U(\varphi \otimes \phi)]), \mathcal{R}_{\mathcal{A}}(P[U(\varphi \otimes \phi)]); P[U(\varphi \otimes \phi)]) = 1.$$

Hence these states are always strongly correlated.

VIII. Examples.

8.1. A particularly interesting class of measurements arises if the coupling is generated by a unitary map of the form

$$(31) \quad U = e^{i\lambda A \otimes B},$$

where A and B are self-adjoint operators in \mathcal{H} and $\mathcal{H}_{\mathcal{A}}$, respectively, and $\lambda \in \mathbf{R}$ is a coupling constant. The operator A is usually taken to represent the (sharp) observable one aims to measure. In order to specify the full measurement scheme and thus the actually measured observable, one needs to choose the pointer observable Z and fix the initial preparation $T_{\mathcal{A}}$ of the apparatus; the measured observable is then given by eq. (2). Using the spectral decomposition of A , $A = \int a E^A(da)$, and denoting

$$(32) \quad T_{\mathcal{A}}^{\lambda a} := e^{i\lambda a B} T_{\mathcal{A}} e^{-i\lambda a B},$$

the final apparatus state, for $T \in \mathcal{S}(\mathcal{H})$, assumes the form

$$(33) \quad \mathcal{R}_{\mathcal{A}}(UT \otimes T_{\mathcal{A}}U^*) = \int \text{tr}[TE^A(da)] T_{\mathcal{A}}^{\lambda a}.$$

Since it is of interest to compare the measured observable E with E^A we assume from the outset that the value space of Z is $(\mathbf{R}, \mathcal{B}(\mathbf{R}))$. In view of the coupling constant λ ($\neq 0$) it is also convenient to introduce a pointer function $f(x) = \lambda^{-1}x$. The observable E measured by the scheme $\langle \mathcal{H}_{\mathcal{A}}, Z, T_{\mathcal{A}}, f, U \rangle$ takes then the following form: for any $X \in \mathcal{B}(\mathbf{R})$,

$$(34) \quad E(X) = \int_{\mathbf{R}} \text{tr}[T_{\mathcal{A}}^{\lambda a} Z(\lambda X)] E^A(da).$$

The structure of the operators $E(X)$ show that in general the measured observable E is not the sharp observable E^A , but a smeared version of it.¹¹ One may ask which choices of Z and $T_{\mathcal{A}}$ would possibly yield $E = E^A$. Obviously, this is the case if and only if for (E^A -almost) all $a \in \mathbf{R}$, $\text{tr}[T_{\mathcal{A}}^{\lambda a} Z(\lambda X)] = \chi_X(a)$, where χ_X is the characteristic function of the set X .

The measurement scheme thus defined is always of the first kind: the measurement outcome probabilities for E are the same both before and after the measurement; for any $T \in \mathcal{S}(\mathcal{H})$ and for all $X \in \mathcal{B}(\mathbf{R})$,

$$(35) \quad \text{tr}[TE(X)] = \text{tr}[UT \otimes T_{\mathcal{A}}U^* E(X) \otimes I].$$

It may also be noticed that the measurement does neither alter the measurement outcome probabilities of E^A , though, as a rule, it is not a measurement of E^A . In fact, if the measurement were an E^A -measurement, it would also be repeatable (3.15) and A would thus have to be discrete, $A = \sum a_i E^A(\{a_i\})$ (3.13). In that case the measurement would also produce all the strong correlations discussed in the previous sections. In general, this is, however, not the case.

Consider next this measurement scheme with a fixed reading scale \mathcal{R} . The pointer observable Z as well as the measured observable E become discretized,

$$(36a) \quad Z^{\mathcal{R}} : i \mapsto Z_i := Z(\lambda X_i),$$

$$(36b) \quad E^{\mathcal{R}} : i \mapsto E_i := E(X_i),$$

and the final component states are

$$\begin{aligned}
(37a) \quad T_S(i, T) &= p_T^E(X_i)^{-1} \mathcal{R}_S(V_i(T)) \\
&= p_T^E(X_i)^{-1} \int \int \mathcal{R}_S(E^A(da) T E^A(da') \otimes Z_i^{1/2} e^{iaB} T_{\mathcal{A}} e^{-ia'B} Z_i^{1/2}) \\
&= p_T^E(X_i)^{-1} \sum t_n \sum L_{kn}^i T L_{kn}^{i*} \\
&\quad \text{with } L_{kn}^i := \int \left\langle \psi_k \mid Z_i^{1/2} \phi_n^{\lambda a} \right\rangle E^A(da) \in \mathcal{L}(\mathcal{H}), \\
&\quad T_{\mathcal{A}} = \sum t_n P[\phi_n] \quad (\text{spectral decomposition}) \\
&\quad \phi_n^{\lambda a} = e^{i\lambda a B} \phi_n, \\
&\quad \{\psi_k\} \subset \mathcal{H}_{\mathcal{A}} \text{ an orthonormal basis,}
\end{aligned}$$

$$\begin{aligned}
(37b) \quad T_{\mathcal{A}}(i, T) &= p_T^E(X_i)^{-1} \mathcal{R}_{\mathcal{A}}(V_i(T)) \\
&= p_T^E(X_i)^{-1} \int \text{tr}[T E^A(da)] Z_i^{1/2} T_{\mathcal{A}}^{\lambda a} Z_i^{1/2},
\end{aligned}$$

(provided that $p_T^E(X_i) \neq 0$). If $E_i^2 = E_i$ for all $i \in \mathbf{I}$, the measurement is repeatable with respect to \mathcal{R} , and

$$(38a) \quad E(X_i) T_S(i, T) = T_S(i, T),$$

$$(38b) \quad T_S(i, T) \cdot T_S(j, T) = O, \quad i \neq j,$$

in which case the implications of theorems 3.11, 5.5, 6.3, 6.4, and 7.2 all hold true. We specify next two instances of the above model, one in which $E_i^2 = E_i$ and another one with $E_i^2 < E_i$.

8.2. Consider a discrete observable $A = \sum a_k E^A(\{a_k\})$, and assume that the set of eigenvalues of A is closed. As the apparatus (or a part of it, called probe) take a particle moving in one-dimensional space, so that $\mathcal{H}_{\mathcal{A}} = L^2(\mathbf{R})$, and couple A with its momentum $P_{\mathcal{A}}$ according to (31). Since the momentum generates translations on the position, it is natural to choose the position $Q_{\mathcal{A}}$ conjugate to $P_{\mathcal{A}}$ as the pointer observable. Assuming that the initial state of \mathcal{A} is a vector state $P[\phi]$, then, in the position representation (for \mathcal{A}) one has $\phi^{\lambda a_k}(x) = \phi(x + \lambda a_k)$, with $\phi^{\lambda a_k} = e^{i\lambda a_k P_{\mathcal{A}}} \phi$. Assuming that the spacing between the eigenvalues a_k is greater than $\frac{\delta}{\lambda}$ and that ϕ is supported in $(-\frac{\delta}{2}, \frac{\delta}{2})$, then the pointer states $\phi^{\lambda a_k}$ are supported in the mutually disjoint sets λI_k , where $I_k = (a_k - \frac{\delta}{2\lambda}, a_k + \frac{\delta}{2\lambda})$. Introducing yet another pointer function g such that $g(I_k) = \{a_k\}$ for each k , and $g((\cup_k I_k)') \subset \{a_k : k = 1, 2, \dots\}'$, one obtains from eq. (34)

$$(39) \quad E(\{a_k\}) = \sum \langle \phi^{\lambda a_i} \mid E^{Q_{\mathcal{A}}}(\lambda I_k) \phi^{\lambda a_i} \rangle E^A(\{a_i\}) = E^A(\{a_k\}),$$

for each k , which shows that the observable measured by this scheme is indeed E^A . The measurement is repeatable, even a Lüders measurement with the state transformer $T \mapsto \mathcal{I}_k(T) = \mathcal{R}_S(V_k(T)) = E^A(\{a_k\}) T E^A(\{a_k\})$, and all the correlations introduced above are strong.

As an elementary quantum optical application, one may consider the measurement of the number observable $N = a^*a$ of a (single-mode) signal field by means of coupling it, via $e^{i\lambda N \otimes b^p}$, with one of the quadrature components $b^p = \frac{i}{\sqrt{2}}(b^* - b)$, say, of another single-mode (probe) field, and using the other quadrature component $b^q = \frac{1}{\sqrt{2}}(b^* + b)$ as the readout observable. With the above choices of the initial probe state ϕ and the pointer functions one obtains a number measurement. It may be noted that neither the beam splitter coupling nor the number-number coupling leads to a sharp number measurement.¹¹

8.3. The second illustration of the above model concerns the case of A being a continuous observable, such as the position of a particle or a quadrature component of a single-mode electromagnetic field. Using the quantum optical nomenclature, we take $A = a^q = \frac{1}{\sqrt{2}}(a^* + a)$, the amplitude quadrature of the (single-mode) signal field with the bosonic annihilation and creation operators a, a^* . For B we take the corresponding quadrature component b^q of a (single-mode) probe field, with the annihilation and creation operators b, b^* . Using the phase quadrature $b^p := \frac{i}{\sqrt{2}}(b^* - b) = \int_{\mathbf{R}} xZ(dx)$ of the probe field as the readout observable, and assuming that the probe field is prepared in a vector state $P[\phi]$ determines the measured observable (34) to be of the form:

$$\begin{aligned}
 E(X) &= \int \int |\hat{\phi}|^2(y - \lambda x) \chi_{\lambda X}(y) dy E^A(dx) \\
 (40) \quad &= \int |\hat{\phi}|^2(y - \lambda a^q) \chi_{\lambda X}(y) dy \\
 &\equiv (e_\lambda * \chi_X)(a^q), \\
 e_\lambda(y) &:= \lambda |\hat{\phi}|^2(-\lambda y),
 \end{aligned}$$

where $e_\lambda * \chi_X$ denotes the convolution of the density function e_λ with the characteristic function of the set X , and $\hat{\phi}$ is the Fourier transform of ϕ .

In the present case the measured observable is the POV measure $E : X \mapsto (e_\lambda * \chi_X)(a^q)$ and not the spectral measure $X \mapsto \chi_X(a^q)$ of a^q ; this is to say that the measured field observable is not the amplitude quadrature a^q but a smearing of it. In fact, if e were replaced by a delta function (concentrated at 0), then (40) would simply give the amplitude quadrature a^q . But this can never occur since the readout observable b^p has no eigenstates, that is, the initial state of the probe field cannot be so chosen that e were a delta function. We observe also that the measurement is not repeatable (since E is not discrete) though still of the first kind. Therefore, the strong correlations are not guaranteed from the outset but need to be studied separately.

Before calculating the observable-correlation produced by the measurement we compare the variance of E with that of a^q in a vector state $P[\varphi]$. Direct application of eq. (2) yields (assuming that the involved quantities are finite)

$$(41) \quad \text{Var}(E, \varphi) = \text{Var}(a^q, \varphi) + \frac{1}{\lambda^2} \text{Var}(b^p, \phi).$$

The initial state $P[\phi]$ of the probe field can be chosen such that $\langle \phi | b^p \phi \rangle = 0$. In this case the measured observable appears, in view of the first moments, as the amplitude quadrature a^q . However, the second moment $\langle \phi | (b^p)^2 \phi \rangle$ never equals 0, meaning that $\text{Var}(E, P[\varphi])$ is strictly greater than $\text{Var}(a^q, P[\varphi])$. However, in the limit of strong coupling, $\lambda \rightarrow \infty$, the measurement noise term $\frac{1}{\lambda^2} \text{Var}(b^p, P[\phi])$ tends to zero. In any case, this shows once more that the actually measured observable is not the amplitude quadrature.

The observable-correlation produced by the measurement is now found to be

$$(42) \quad \rho(E, Z^f; P[U(\varphi \otimes \phi)]) = \frac{\text{Var}(a^q, P[\varphi])}{\text{Var}(E, P[\varphi])},$$

a quantity always strictly less than 1. The measurement, though of the first kind, does never lead to strong observable-correlation. Yet,

$$(43) \quad \lim_{\lambda \rightarrow \infty} \rho(E, Z^f; P[U(\varphi \otimes \phi)]) = 1.$$

In order to discuss the value- and state-correlations produced by the measurement scheme one needs to introduce a reading scale \mathcal{R} . The discrete observable $E^{\mathcal{R}} : i \mapsto E_i$ thus measured is

$$(44) \quad E_i = (e_\lambda * \chi_{X_i})(a^q),$$

whereas the final component states (37a) are of the form:

$$(45) \quad T_{\mathcal{S}}(i, P[\varphi]) = \langle \varphi | E_i \varphi \rangle^{-1} \int_{\lambda X_i} L_y P[\varphi] L_y^* dy$$

with $L_y := \hat{\phi}(y - \lambda a^q)$.

Neither the eigenvalue condition (38a) nor the orthogonality condition (38b) can be satisfied for all initial vector states of the signal field. Therefore the strong value and state-correlations cannot be inferred by using theorems 6.3 and 7.2. Still the value-correlation is always strong: $\rho(E_i, Z_i; P[U(\varphi \otimes \phi)]) = 1$ for all i and for any $P[\varphi]$ for which $\langle \varphi | E_i \varphi \rangle \neq 0$. Indeed, due to the commutativity of the operators L_y of eq. (45) with E_i , ϵ_{12} of eq. (28a) equals $\langle \varphi | E_i^2 \varphi \rangle$; furthermore the first kind property of the measurement and the sharp pointer yield for (28b-d): $\epsilon_1 = \epsilon_2 = \langle \varphi | E_i \varphi \rangle$, and $\sigma_1^2 = \sigma_2^2 = \text{Var}(E_i, P[\varphi])$. Therefore $\epsilon_{12} - \epsilon_1 \epsilon_2 = \sigma_1 \sigma_2$, so that $\rho(E_i, Z_i; P[U(\varphi \otimes \phi)]) = \text{Var}(E_i, P[\varphi]) / \text{Var}(E_i, P[\varphi]) = 1$. Finally, a direct computation of the state-correlation $\rho(T_{\mathcal{S}}(i, P[\varphi]), T_{\mathcal{A}}(i, P[\varphi]); P[U(\varphi \otimes \phi)])$ shows that this number is not, in general, equal to one.

IX. Conclusion.

In this paper we have investigated possible properties of the final component states of the object system and the apparatus (or probe) arrived at in a quantum measurement, properties which must be required if the occurrence of definite measurement outcomes is to be understood as the conjunction of pointer value definiteness (PVD), pointer mixture property (PM), plus the ignorance interpretation

for the final reduced apparatus state. According to Theorem 3.11, the properties (PVD) and (PM) are ensured if the final component states of the object system are mutually orthogonal. Considering initial states of \mathcal{S} which are vector states, this latter condition is also necessary for (PM) in the case of a unitary measurement \mathcal{M}_U , where (PVD) is automatically fulfilled since the pointer is a sharp observable. The orthogonality of the states $T_{\mathcal{S}}(i, T)$ is not always guaranteed.¹²

Next we have considered conditions for strong correlations between observables, their values, or between the final component states of \mathcal{S} and \mathcal{A} . It turns out that repeatable measurements give strong observable- as well as strong value-correlations (Theorems 5.5, 6.3, 6.4). Furthermore, strong observable-correlation for finite reading scales entails repeatability and thus the orthogonality of the states $T_{\mathcal{S}}(i, T)$ and hence (PVD), via 3.11. On the other hand, strong value-correlation may occur independently of (PVD). Finally, strong state-correlation may occur under more general circumstances than the other correlations since it is independent of the repeatability property, but its implying the orthogonality of the final component states of \mathcal{S} may be limited to the case where these states are vector states. However in that case, and for a unitary measurement \mathcal{M}_U , strong state-correlation is equivalent to the said orthogonality and thus to the pointer mixture condition.

In conclusion, we wish to emphasize that our investigation provides an illustration of how interpretational demands entail formal constraints on measurements that may or may not be fulfilled in a concrete case. These formal features have thus to be made explicit if the consistency of an interpretation is to be demonstrated. With these findings we believe to have settled the questions left open in previous work.²

Acknowledgements. Part of this work was carried out while one author (PB) was Visiting Research Scholar at the Lyman Laboratory of Physics, Harvard University, Cambridge, MA. This visit was funded by means of a Feodor Lynen Fellowship extended to him by the Alexander von Humboldt-Foundation, Bonn. Support and hospitality of Harvard University are gratefully acknowledged.

References.

1. P. Busch, P. Lahti, P. Mittelstaedt, *The Quantum Theory of Measurement*, **LNP m2** (Springer-Verlag, Berlin, 1991). Second revised edition forthcoming.
2. G. Cassinelli, P. Lahti, *Nuovo Cimento B* **108**, 45 (1993).
3. D. F. Walls, G. J. Milburn, *Quantum Optics* (Springer-Verlag, Berlin, 1994).
4. M. Ozawa, *J. Math. Phys.* **25**, 79 (1984).
5. G. Cassinelli, N. Zanghi, *Nuovo Cimento B* **73**, 237 (1983).
6. G. Cassinelli, N. Zanghi, *Nuovo Cimento B* **79**, 141 (1984).
7. A. Łuczak, Instruments on von Neumann algebras, Institute of Mathematics, Łódź University, Poland, 1986.
8. P. Lahti, P. Busch, P. Mittelstaedt, *J. Math. Phys.* **34**, 2770 (1991).
9. P. Halmos, *Measure Theory*, **GTM 18** (Springer-Verlag, New York, 1988).
10. S. K. Berberian *Notes on Spectral Theory*, (D. Van Nostrand Company, Inc.

Princeton, 1966).

11. P. Busch, M. Grabowski, P. Lahti, *Operational Quantum Physics*, **LNP m31** (Springer-Verlag, Berlin, 1995).
12. E. Beltrametti, G. Cassinelli, P. Lahti, *J. Math. Phys.* **31** (1990) 91.